

Comments on black holes I: The possibility of complementarity

Samir D. Mathur and David Turton

Department of Physics,
The Ohio State University,
Columbus, OH 43210, USA

mathur.16@osu.edu
turton.7@osu.edu

Abstract

We comment on a recent paper of Almheiri, Marolf, Polchinski and Sully who argue against black hole complementarity based on the claim that an infalling observer ‘burns’ as he approaches the horizon. We show that in fact measurements made by an infalling observer outside the horizon are statistically identical for the cases of vacuum at the horizon and radiation emerging from a stretched horizon. This forces us to follow the dynamics all the way to the horizon, where we need to know the details of Planck scale physics. We note that in string theory the fuzzball structure of microstates does not give any place to ‘continue through’ this Planck regime. AMPS argue that interactions near the horizon preclude traditional complementarity. But the conjecture of ‘fuzzball complementarity’ works in the opposite way: the infalling quantum is absorbed by the fuzzball surface, and it is the resulting dynamics that is conjectured to admit a complementary description.

Contents

1	Introduction	1
2	The information paradox and the fuzzball proposal	3
3	Traditional complementarity vs Fuzzball complementarity	6
4	Limits on measurements made outside the horizon	13
5	The AMPS argument	18
6	Discussion	23
A	Appendices	26

1 Introduction

The quantum theory of black holes has proven to be rich territory for the exploration of the most fundamental laws of physics. The discoveries of black hole entropy [1], and Hawking radiation [2] provided deep links between gravity and thermodynamics, while raising a serious problem in the form of the information paradox. One suggestion that arose in this context was the notion of black hole complementarity [3]. String theory provides a microscopic explanation for the entropy of black holes [4], and the fuzzball structure of microstates provides a solution to the information paradox [5, 6, 7, 8].

Recently there have appeared several papers discussing the relations between the information paradox, entanglement theorems, complementarity and other issues involving the quantum theories of black holes [9, 10]. Since there are several interrelated issues in the area of black holes, we have split our discussion into a set of papers, each addressing a different question. In this article we address some of the arguments used in the paper of Almheiri, Marolf, Polchinski and Sully (AMPS) [9] and show that they are not valid as they stand.

We note that the fuzzball program provides a consistent picture of all issues in the quantum dynamics of black holes (see [11] for reviews). We will keep this fact at the back of our mind, since in many cases the fuzzball description provides us an explicit model to judge the validity of abstract arguments.

We begin with some definitions and basic facts about black holes and the information paradox. We then make two observations:

- (a) It is often assumed that if an infalling observer ‘hits something’ at the horizon,

then there cannot be a ‘complementary’ description where he goes through. While traditional complementarity may have this feature, the kind of complementarity suggested by fuzzballs is different. We use a toy example provided by AdS/CFT duality to observe that in one description an infalling quantum ‘breaks up’, while in another description it continues its trajectory unscathed. We note that the case of the black hole is a little different from the AdS/CFT case, and explain how complementarity can arise for quanta falling into the hole with energy $E \gg kT$.

(b) One might think that an observer falling into the traditional black hole sees nothing as he falls up to the horizon, but an observer falling towards a body radiating ‘real quanta’ from a stretched horizon would get ‘burnt’ by the highly energetic photons encountered close to this horizon. We show that observations of Hawking quanta made outside the horizon actually yield similar results in both cases. Switching off a detector before crossing the horizon of a traditional black hole creates excitations from vacuum fluctuations, and these excitations have the same spectrum as excitations created by ‘real quanta’ from a stretched horizon.

We then address the argument made in AMPS [9]. In brief outline, the AMPS argument goes as follows:

- (i) If Hawking evaporation is unitary, then the state near the horizon is not the vacuum in an infalling observer’s frame, but involves high-energy excitations.
- (ii) If there are high-energy excitations near the horizon, then an infalling observer will measure physical high energy quanta emerging from the black hole, and get burnt.
- (iii) If the observer gets burnt, then we cannot have any complementary description where he falls through without noticing anything at the horizon.

From points (a) and (b) above, we find that the AMPS gedanken experiment does not lead to the conclusions they suggest. If one wishes to avoid planck scale physics, then one should restrict to measurements made outside the stretched horizon. For such measurements point (b) shows that an infalling observer will see the traditional black hole and a radiating stretched horizon as statistically similar systems. The underlying reason for this equivalence is that there is too little time for him to detect the Hawking quanta before he reaches the horizon. More importantly, point (a) shows that even if the infalling observer were to hit the stretched horizon violently, this fact would not by itself invalidate the possibility of complementarity; in fact it is this very interaction that is expected to admit a complementary description.

In the discussion (Section 6) we summarize the essential physics involved in the conjecture of fuzzball complementarity to show precisely where the AMPS argument goes wrong.

The reader who is already familiar with fuzzballs and the conjecture of fuzzball complementarity may skip directly to Section 4.

2 The information paradox and the fuzzball proposal

In this section we review the resolution of the information paradox through the fuzzball construction in string theory. Though the later arguments will be more abstract, the steps below will help us decide the validity of these arguments.

(a) The traditional black hole

The information paradox arises from the way Hawking radiation is emitted from the *traditional black hole*. We define the traditional black hole as follows. There is a horizon, and a neighbourhood of the horizon with the following property. One can choose good slices in this neighbourhood, and in these good coordinates physics is ‘normal’. Here ‘normal’ physics means exactly what we mean by normal physics in the lab: evolution of long wavelength modes ($\lambda \gg l_p$) is given by local quantum field theory on curved space, with corrections controlled by a small parameter ϵ . These corrections can come from any quantum gravity effect, local or nonlocal, and all we require is that $\epsilon \rightarrow 0$ as $M \rightarrow \infty$, where M is the mass of the black hole.

(b) The information paradox

The traditional black hole arose from a study of gravitational collapse that leads to the Schwarzschild metric

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \frac{dr^2}{1 - \frac{2M}{r}} + r^2 d\Omega_2^2 \quad (2.1)$$

If we use semiclassical gravity to follow the evolution of quantum modes during the collapse, we get the traditional black hole. We have the a vacuum region around the horizon which indeed gives ‘lab’ physics in a good slicing (i.e., in Kruskal coordinates). Evolution of vacuum modes at this horizon leads to entangled pairs being created, with one member of the pair staying in the black hole and the other escaping to infinity as Hawking radiation. The entangled pair can be modeled for simplicity by [12]¹

$$|\psi\rangle_{pair} = \frac{1}{\sqrt{2}} (|0\rangle_{in}|0\rangle_{out} + |1\rangle_{in}|1\rangle_{out}) \quad (2.2)$$

The entanglement between the inside and outside grows by $\ln 2$ with each emission. Near the endpoint of evaporation this would leave just two possibilities: information loss or a

¹Further analysis of such ‘bit models’ can be found in [13, 14, 15].

remnant [2, 12]. Both of these look unsatisfactory; we would like a pure state of Hawking radiation carrying all the information of the black hole.

(c) The theorem controlling small corrections

The problem would be resolved if gravitational collapse led to a state other than the traditional black hole. But the traditional black hole solution appeared to admit no deformations, leading to the phrase ‘black holes have no hair’. Exactly the same problem holds for black holes in AdS. Thus AdS/CFT duality cannot by itself help to resolve the problem (for a detailed discussion of this issue, see [12, 16]).

This situation led many string theorists to the following belief. Hawking computed the pair creation at leading order, but there can always be small quantum gravity corrections to the wavefunction (2.2)

$$|\psi\rangle_{pair} = \frac{1}{\sqrt{2}} (|0\rangle_{in}|0\rangle_{out} + |1\rangle_{in}|1\rangle_{out}) + \epsilon \frac{1}{\sqrt{2}} (|0\rangle_{in}|0\rangle_{out} - |1\rangle_{in}|1\rangle_{out}) \quad (2.3)$$

where we have added a small amount of an orthogonal state for the pair. The correction ϵ for each pair must be small since the horizon geometry is smooth, but the number of emitted quanta is large ($\sim (M/m_p)^2$), and the net effect of the small corrections may accumulate in such a way that the overall state of the radiation would not be entangled with the black hole.

But in [12] it was shown that this hope is false; the change in entanglement δS_{ent} , compared to the entanglement S_{ent} of the leading-order Hawking process, is bounded by

$$\frac{\delta S_{ent}}{S_{ent}} < 2\epsilon. \quad (2.4)$$

This inequality is the essential reason why the Hawking argument has proved so robust over the years – no small corrections can save the situation. We will make use of (2.4) many times; many arguments in the other papers we discuss are also based on this inequality.

(d) The fuzzball structure of microstates

In [17] it was found that a bound state in string theory *grows* in size with the number of branes in the bound state and with the coupling, so that its wavefunctional is always spread over a radius which is order the Schwarzschild radius. This growth in size is a very stringy effect; it arises from the phenomenon of ‘fractionation’ [18] which uses the extended nature of fundamental objects in the theory. Such horizon sized wavefunctionals are termed ‘fuzzballs’.

The size of fuzzball states is estimated by using the entropy of brane bound states, together with the physics of fractionation. Thus this size estimate involves *all* the states

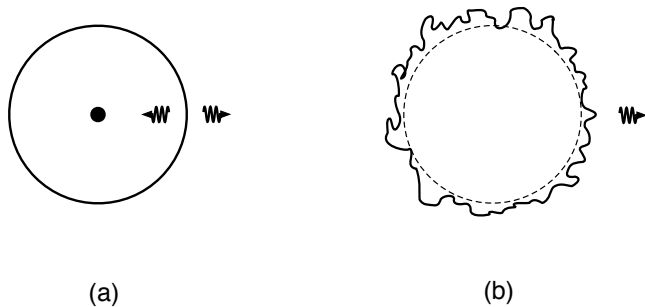


Figure 1: (a) The traditional black hole; small corrections at the horizon *cannot* get information out in the Hawking radiation. (b) The fuzzball picture of black hole microstates; spacetime ends in stringy theory sources just before the horizon is reached.

of the black hole. To study the properties of fuzzballs further, it is useful to look at states where we place ‘many quanta in the same mode’. This is analogous to black body radiation, where placing a large number N of quanta in the same harmonic gives a laser beam, with quantum fluctuations suppressed as $\sim \frac{1}{N}$.² One finds that the fuzzballs generate a spacetime that resembles the traditional black hole far away from the horizon, but which ends in a set of string theory sources before reaching the horizon [5, 6, 7]. This is pictured in fig.1.

(e) Resolution of the paradox

Given the existence of fuzzballs, the information paradox is resolved as follows. Fuzzballs do not radiate by pair creation from an ‘information-free horizon’; instead the radiation emerges from the surface of the fuzzball and carries information just like any normal body. This radiation has been explicitly worked out for simple fuzzballs; the rate of radiation agrees exactly with the Hawking emission rate expected for those fuzzballs but the details of the fuzzball state are seen to be imprinted in the spectrum of emitted quanta [19].

If we start with a collapsing shell, then its wavefunction spreads over the enormous phase space of fuzzball states [20], and then these fuzzball states radiate like any other warm body. The time for this spread can be estimated to be much smaller than the

²This study of low fluctuation states has led some to be confused about the nature of fuzzballs. They ask: are fuzzballs just solutions to supergravity or do they involve stringy degrees of freedom? As can be seen from the above discussion, there is no fundamental classical/quantum divide between states; all we can do is look at states with small or large fluctuations. In particular the nonextremal states studied in [22] using the pp-wave technique were given in terms of strings placed in a fuzzball geometry. The correct question is not; ‘how messy is the fuzzball’; the only relevant question is ‘do we get a traditional black hole (with ‘lab physics’ around a horizon) or do we not’. The only feature common to all fuzzballs is that we never form a traditional horizon.

Hawking evaporation time [21]

$$t_{\text{fuzzball}} \ll t_{\text{hawking}} \quad (2.5)$$

This solves the information paradox.

3 Traditional complementarity vs Fuzzball complementarity

In this section we will explain what we mean by having a ‘complementary description’. We start by giving a toy example: the case of AdS/CFT duality [23]. We briefly recall the traditional notion of complementarity, and then turn to how complementarity is conjectured to arise in the fuzzball description of microstates. This ‘fuzzball complementarity’ has things in common to the toy example of AdS/CFT duality, but also differs from it in a crucial way.

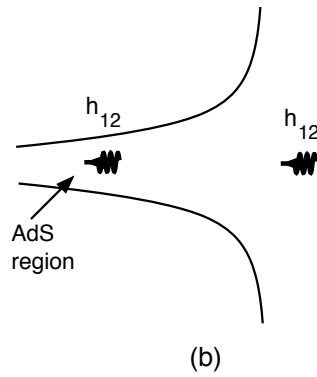
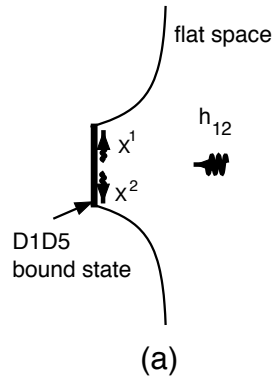
3.1 Toy example of complementarity: AdS/CFT duality

We start with an example that illustrates what we mean by having a complementary description. In this example an infalling quantum will encounter some degrees of freedom and appear to ‘go splat’; i.e. get ‘destroyed’. Yet there will be an alternative description where it continues unscathed. When a description of the latter kind exists, we will say that we have a ‘complementary’ description of the degrees of freedom in the former description.

Consider IIB string theory compactified on $S^1 \times T^4$. Let y be the coordinate along S^1 and z_1, \dots, z_4 be the coordinates on T^4 . We consider a bound state of n_1 D1 branes wrapped on S^1 and n_5 D5 branes wrapped on $S^1 \times T^4$. This bound state is depicted in fig.2(a), where the direction along the branes is the S^1 .

A graviton with both indices on the T^4 is a scalar in the remaining dimensions. Consider in particular the graviton h_{12} , arriving at the brane bound state as shown in fig.2(a). On hitting the brane bound state, the energy of the graviton gets converted to vibrations of the branes (open strings);³ a vibration polarized in the direction X^1 moves up along the S^1 and a vibration polarized in the direction X^2 moves down the S^1 [24].

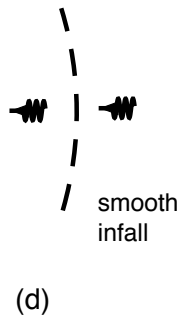
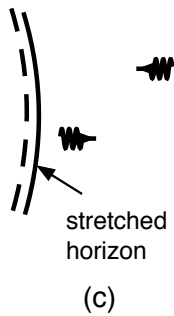
³The actual evolution on the branes is more complicated when we consider interactions in the CFT, but this simple picture illustrates the point we wish to make. Note that we are considering the gravity description at weak coupling, and so the CFT description is at strong coupling. But the important fact is that there are *two* descriptions at the same coupling; one using strongly interacting spin 1 gauge degrees of freedom, and one using the spin 2 graviton and higher closed string modes. In the former description the incoming h_{12} appears to break up into pieces, while in the latter it remains intact.



AdS/CFT duality (Toy example)

(a) A graviton coming in from flat infinity hits the D1D5 branes and breaks up.

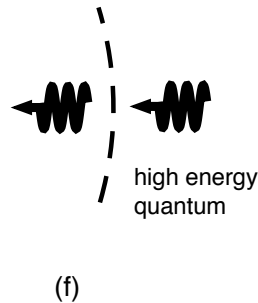
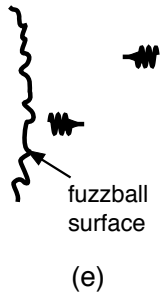
(b) In a complementary description, the graviton passes unchanged into an AdS region.



Traditional complementarity

(c) A stretched horizon absorbs and emits quanta in the frame of the observer at infinity.

(d) An infalling quantum passes smoothly through the horizon in the frame of an infalling observer.



Fuzzball complementarity

(e) The spacetime ends in string theory sources outside $r=2M$, with no horizon. Unitary Hawking radiation ($E \sim kT$) is emitted from the fuzzball surface.

(f) $E \gg kT$ quanta excite collective dynamics of the fuzzball. This dynamics has an approximate complementary description in terms of smooth infall.

Figure 2: AdS/CFT duality, traditional complementarity and fuzzball complementarity.

One may say that the graviton has ‘gone splat’ on hitting the branes, to such an extent that it has split into two parts, X^1 and X^2 . These two products obtained after impact certainly do not look like the single graviton h_{12} that was arriving towards the brane bound state. But as we well know, there is an alternative description of this physics where we replace the brane bound state by an AdS region. In the latter description, the graviton h_{12} falls smoothly into the AdS region, remaining as a single entity h_{12} (fig.2(b)). We can call this latter description a ‘complementary’ description of the interaction with the D1D5 branes. Now we can ask our question: when the incoming graviton broke up on the D1D5 brane bound state, did it go ‘splat’ or not?

To better understand how to interpret this situation, we look at a more detailed example where we start with *two* gravitons, h_{12} and h_{34} , separated by a distance D . We can think of this pair of gravitons as being an ‘object’; if the separation of the gravitons is increased or decreased, we can say that the object has ‘been damaged’ and ‘feels pain’.

The evolution of these gravitons proceeds as follows [5]. First h_{12} hits the D1D5 bound state, and changes to excitations X^1, X^2 which travel at the speed of light in opposite directions along y . At a later time h_{34} hits the bound state and changes to vibrations X^3, X^4 , again separating at the speed of light. But the separation D between the initial gravitons can be recovered from the open string excitations. Let y_i be the location along the S^1 of the excitation X^i , for $i = 1, 2, 3, 4$. We have

$$y_1 = t, \quad y_2 = -t, \quad y_3 = t - D, \quad y_4 = -t + D \quad (3.1)$$

so the value of D is encoded in the vibrations X^i as

$$D = \frac{1}{2}[(y_1 - y_2) - (y_3 - y_4)] \quad (3.2)$$

In the dual gravity description, the two gravitons fall smoothly into the AdS, maintaining their separation D and thus showing no indication of ‘damage’ or ‘pain’. But given that the CFT description is a faithful copy of the gravity description, and that we can recover the same value D from the CFT, it looks correct to say that there is no damage or pain felt in the brane description either.

By contrast, when we throw an object onto a normal concrete wall, we do not expect to find a complementary description. Let us analyze what was special about the D1D5 brane case which did allow for complementarity.

In the D1D5 example, the Hilbert space of the incoming gravitons mapped faithfully into the Hilbert space of vibrations of the branes. That is, if we write the eigenstates of the incoming graviton as $|\psi_E\rangle$ and the eigenstates of the D1D5 system as $|\tilde{\psi}_E\rangle$, then we find

$$\int dE C(E) |\psi_E\rangle \rightarrow \int dE C(E) |\tilde{\psi}_E\rangle \quad (3.3)$$

The *nature* of the excitations changed completely - they changed from being gravitons to being vibrations of branes - but this is *not* important. What is important is that the amplitude for a given energy remained the same (or approximately the same). A important input for getting a relation like (3.3) is that the D1D5 bound state had a very closely spaced set of energy levels. This high density of levels leads to a ‘fermi-golden-rule’ absorption of the graviton, and in such an absorption each incoming energy level E_k transfers its amplitude to energy levels \tilde{E}_k that are very close to E_k . (In [25] the absorption of the graviton onto the brane bound state was computed by such a fermi-golden rule process.)

What *does* cause ‘damage’ or ‘pain’ is the situation where the levels available in the absorbing system are not sufficiently continuous. In this situation we will find in general

$$\int dE C(E) |\psi_E\rangle \rightarrow \int dE C'(E) |\tilde{\psi}_E\rangle, \quad C(E) \neq C'(E) \quad (3.4)$$

In particular, a concrete wall will not have the same energy levels as the object hitting it, and so the incoming object will not be mapped faithfully into excitations of the concrete wall. In this situation we do not expect a complementary description of the impact.

To summarize, we cannot just say: ‘If we go ‘splat’ on hitting some degrees of freedom, then we cannot have complementarity’. The impact transfers excitation energy to the degrees of freedom that are encountered. To know if we can have a complementary description we have to ask if the Hilbert space of the infalling object maps faithfully into a subspace of the Hilbert space of the encountered degrees of freedom.

3.2 Traditional complementarity

In the early works on black hole complementarity [3], the physics that was proposed is depicted in fig.2(c),(d). It was assumed that we can place a ‘stretched horizon’ just outside $r = 2M$, and that incoming quanta could be taken to interact with degrees of freedom on this stretched horizon. In the complementary description, we have just the smooth infall through the horizon.

The problem with this proposal is discovered when we ask for the physical origin of the degrees of freedom on the stretched horizon. It was argued that since the Schwarzschild coordinates break down at $r = 2M$, there will be violent fluctuations of the gravitational degrees of freedom as we approach $r = 2M$. It was further argued that these violent fluctuations are indicative of the fact that physics outside the horizon is self-consistent, and the stretched horizon provides the natural boundary beyond which we need not look.

Such an argument is, however, unsatisfactory. The breakdown of Schwarzschild coordinates means that we should use better coordinates, not that we are entitled to assume

new physics. But there is an even more serious difficulty with this proposal, which we can see by returning to our basic question: how does the information paradox get resolved? There is a ‘smooth slicing’ of the geometry where we see the creation of entangled pairs (2.2). The defenders of traditional complementarity argued that the inner and outer parts of the horizon should not be considered in the same Hilbert space, since an observer who falls in has strong limitations on how he can communicate with the outside; thus the state (2.2) makes no sense. But no mechanism was proposed to implement such a drastic change to normal physics. The skeptics of complementarity simply noted that there *is* a good slicing of the geometry which we should use to do physics at the horizon, and with this slicing there appears to be no reason to not have a single Hilbert space that includes both the inner and outer parts of the horizon.

For these reasons, the traditional picture of complementarity remained an unresolved issue. It is important to note the difference between the traditional black hole case and the example of AdS/CFT that we presented in Section 3.1. In the AdS/CFT example of fig.2(a),(b), the boundary where we get a complementary description is not a horizon, and there is no particle creation there. Thus we do not have the information problem. But in the case of a black hole there is no way to stop the creation of entangled pairs in any picture where a smooth horizon is assumed, and then we cannot escape the information paradox. As we will see now, the way complementarity can arise with the fuzzball picture in string theory is somewhat different, and needs us to recognize that real degrees of freedom appear at the location of the horizon.

3.3 The proposal of fuzzball complementarity

With the explicit construction of black hole microstates in string theory (fuzzballs) we find that things work out differently from the traditional picture of complementarity. The general idea of ‘fuzzball complementarity’ is developed in [26, 13, 27, 28]. The notion of making spacetime by entanglement [29, 30, 31] is very useful in this approach. Here we just give an outline of how things work:

(a) Complementarity does *not* arise because of a choice of coordinate frame (Schwarzschild vs Kruskal). Instead, the construction of microstates is fully covariant.

(b) In the traditional black hole we have *vacuum* around the horizon. But in string theory, spacetime has a ‘boundary’ where it ends with in a set of string theory sources just outside $r = 2M$, before the horizon is reached. The details of these sources encode the choice of microstate.

(c) Hawking radiation arises as quanta radiated from the details of microstate structure near the boundary. For simple microstates this radiation has been explicitly com-

puted, and it arises from ‘ergoregion emission’ [19] near the boundary. The details of the ergoregion structure depend on the choice of microstate.

(d) Since we have ‘real’ degrees of freedom at the horizon, the $E \sim kT$ quanta radiated from the microstate is able to carry out the information of the microstate. We *cannot* have a complementary picture where we replace the physics of such quanta by the vacuum physics seen at the horizon of the traditional black hole. In this way our complementarity differs from traditional complementarity. What we have to do is make a distinction between $E \sim kT$ quanta (relevant for the information problem) and $E \gg kT$ quanta (relevant for the ‘infall problem’ of heavy observers). Only the latter quanta will have a complementary description.

(e) The complementarity conjecture is now the following (fig.2(e),(f)). When $E \gg kT$ quanta impact the fuzzball surface, the resulting dynamics can be reproduced to a first approximation by the geometry of the black hole interior, for times of order crossing time (i.e. before the quantum hits the singularity). This description emerges from the fuzzball dynamics as follows. The $E \gg kT$ quantum excites ‘collective modes’ of the fuzzball. In a first approximation the evolution of these modes is insensitive to the precise choice of fuzzball microstate (assuming we have taken a generic microstate). The evolution of these collective modes in this leading approximation is to be encoded in a ‘complementary’ description. Thus let the initial state of the hole have mass M and be the linear combination of fuzzball states $\sum_i C_i |F_i\rangle$. When a quantum of energy $E \gg kT$ falls onto the fuzzball surface, the wavefunction of the fuzzball shifts to a combination $\sum_j C'_j F'_j$ over the fuzzball states with mass $M + E$:

$$\sum_i C_i |F_i\rangle \rightarrow \sum_j C'_j F'_j \quad (3.5)$$

If $E \gg kT$ then the number of coefficients C'_j is much larger than the number of C_i . The leading order evolution of the coefficients C'_j is to be captured by the complementary description.

(f) We can now see the similarities and differences with the toy example of AdS/CFT duality discussed in Section 3.1:

(i) The D1D5 brane degrees of freedom correspond to degrees of freedom at the ‘boundary’ of the fuzzball microstate.

(ii) The D1D5 branes were taken to be in their ground state,⁴ while the fuzzball structure differs microscopically from state to state. Thus we get only approximate

⁴We can take excited states of the D1D5 branes, but in AdS/CFT duality we take these to be low energy excitations, and their effect in the dual gravitational description will occur near $r = 0$, not near the place where the CFT is placed

complementarity in the black hole case, by looking at energies $E \gg kT$ where the detail of the fuzzball microstate become irrelevant.

(iii) In the AdS/CFT case the complementary description was possible because of the closely spaced levels of the D1D5 brane system. In the black hole case we again have a close spacing of levels because of guaranteed by the large number $Exp[S_{bek}]$ of fuzzball microstates.

3.4 Summary

To summarize, we have observed the following:

(a) In our toy example of AdS/CFT duality, we have a brane description, where an incoming quantum appears to hit some degrees of freedom violently and ‘break up’. In a ‘complementary description’, the incoming quantum smoothly through into an AdS region. There is no radiation from the AdS boundary itself, so there is no creation of entangled pairs at that location.

(ii) In traditional complementarity, one argues that there are two equivalent descriptions, a fact allowed by the limitations on communication between observers inside and outside the hole. In one description (that of the outside observer) incoming quanta are reflected back as Hawking radiation from a stretched horizon, while in another description (that for an infalling observer) the horizon is a smooth place. Since there is a horizon, there is a creation of entangled pairs (2.2) in a smooth slicing at that location, and there is no clear mechanism to remove this entanglement.

(iii) In fuzzball complementarity, there are real degrees of freedom at the horizon which arise from the fact for each black hole microstate, the compact directions pinch off in a mess of string sources and spacetime ends before we reach $r = 2M$. The details of this ‘fuzzball’ differs from microstate to microstate; there is no Hawking type creation of entangled pairs and the radiation from the fuzzball surface can be explicitly seen carry information of the microstate. Since the fuzzball surface differs from microstate to microstate, complementarity can only be obtained in an approximation where the effect of these differences is small. The conjecture is when $E \gg kT$ quanta impact the fuzzball, they excite collective modes that are relatively insensitive to the precise choice of microstate; the evolution of these modes (3.5) can be approximated by evolution in a spacetime that mimics the black hole interior.

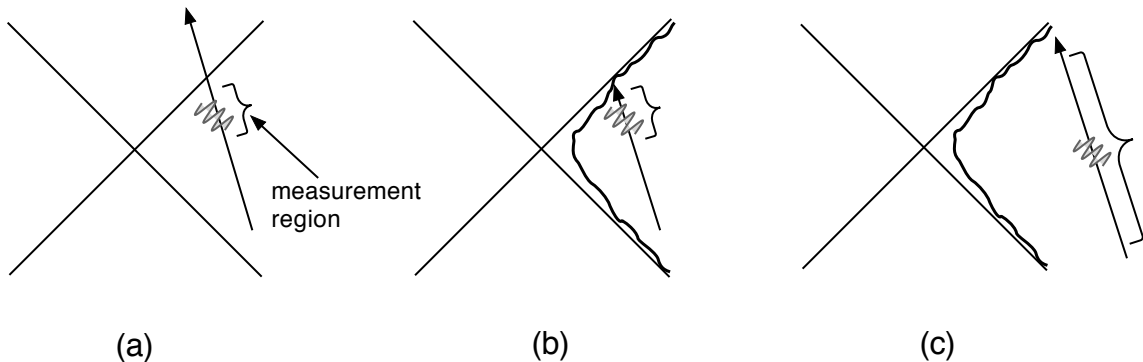


Figure 3: (a) An inertial detector in Minkowski space, making a measurement using only the indicated part of its trajectory. Vacuum fluctuations excite the detector. (b) A similar detection, but for case of a warm body radiating into the right Rindler wedge. The wavelength of quanta is of the same order as the distance from the horizon. (c) Radiation from a ‘hot’ body, where the wavelength is much shorter than the distance from the horizon.

4 Limits on measurements made outside the horizon

In this section we address the following question. If we measure the radiation outside a black hole, then can we tell the difference between a traditional black hole and an object that radiates unitarily at the same temperature T from a surface just outside $2M$?

The measurements we are interested in are close to the horizon ($|r - 2M| \ll 2M$), so we can consider the near horizon geometry depicted in fig.3. In fig.3(a) we have the traditional black hole, which has vacuum around the horizon, so the near horizon region looks like Minkowski space when seen in Kruskal coordinates. In fig.3(b) we have a warm surface placed just outside $r = 2M$ (indicated by the jagged line), and this surface is assumed to radiate quanta at the temperature T of the black hole.

At first it may appear that the case of fig.3(b) has real radiation that can ‘burn’, while there is no real radiation in fig.3(a). We *can* see quanta in Minkowski spacetime by taking a detector that accelerates. But our interest is in freely falling observers, which are indicated by the straight line trajectory in fig.3(a). One may expect that a detector moving in straight line in Minkowski space should not detect any quanta. But the situation we have is a little special. We are asking if we can distinguish the physical situations of fig.3(a) and fig.3(b) by observations *outside the horizon*. Thus a detector trying to make a measurement would have to do this task by using only a section of its trajectory like that indicated in fig.3(a).

But if we place conditions on how long a detector has to make a measurement, then

we run into the problem that we pick up vacuum fluctuations. We discuss the scales involved in the problem in Section 4.1. Suppose we are considering radiation at the Hawking temperature T . The wavelength of these quanta at a distance d from the horizon is $d \sim \lambda$. The infalling detector trying to measure such quanta a limited time to make this measurement, and we argue that this available time is less than the time required to make the desired measurement.

In Section 4.2 we note that the above estimates reflect a general fact: for generic states of the radiating body in fig.3(b), observations of radiation do not appear statistically different from the vacuum fluctuations picked up by the detector of fig.3(a). The arguments we give are very basic to the theory of particle detection, and are implicit in many treatments of Rindler space.⁵

4.1 Time needed for detector response

Let us examine what kind of quanta a detector can actually pick up in a measurement process. In Appendix A we show that if we wish to measure a quantum of wavelength λ , then we need a proper time $\gtrsim \lambda$ to elapse along the detector trajectory:

$$\Delta\tau_{needed} \gtrsim \lambda \quad (4.1)$$

In Fig.3 we note two different possibilities for the location of the quantum of wavelength λ . In fig.3(a),(b) the quantum is at a distance $d \sim \lambda$ from the black hole surface. In fig.3(c) the quantum is at a distance $d \gg \lambda$ from the black hole surface. In Appendix B we show that the Hawking quanta radiated from the black hole surface are of the former type; the typical wavelength found at a distance λ from the horizon is $\sim \lambda$ itself:

$$\lambda \sim d \quad (4.2)$$

We now begin to see the source of difficulty in catching high energy Hawking quanta: we are already very close to the horizon when we encounter them, and then we may have too little time left to interact with them. Before proceeding, there is one effect that we must take into account. Because the detector is infalling, it sees the outgoing quantum as being Lorentz contracted; thus the wavelength of the quantum appears shorter than the distance d measured along a $t = \text{const}$ slice. We take a local Lorentz frame oriented along the Schwarzschild t, r directions, and let the proper velocity of the detector in this frame be

$$U^{\hat{t}} = \cosh \alpha, \quad U^{\hat{r}} = -\sinh \alpha \quad (4.3)$$

⁵We also thank Bill Unruh for an earlier conversation about detectors in Minkowski spacetime.

Then, as shown in Appendix B, the effective wavelength of the Hawking quanta encountered by the infalling detector is

$$\lambda_{eff} \sim de^{-\alpha} \quad (4.4)$$

Now we consider the proper time available to an infalling detector to measure the Hawking quantum; this detection must be made between the time the detector is at a distance $\sim d$ from the horizon and the time it falls through the horizon. In Appendix C we show that for a detector falling in from far outside the horizon, this proper time is

$$\Delta\tau_{available} < de^{-\alpha} \quad (4.5)$$

Putting together (4.1), (4.2) and (4.5) we get

$$\Delta\tau_{available} < \Delta\tau_{needed} \quad (4.6)$$

so we conclude that an infalling detector cannot reliably pick up Hawking quanta being radiated from a black hole surface. We now turn to comparing the behavior of detectors in the situations of fig.3(a) and fig.3(b).

4.2 Detectors in Rindler space and detectors near warm bodies

Let us consider the following question. We look at the situation of fig.3(a), where we have an inertial detector in empty Minkowski space, but the detection is required to be made before the detector crosses the Rindler horizon. We can therefore capture our physics by using Rindler coordinates covering the right Rindler wedge

$$t_M = r_R \sinh t_R, \quad x_M = r_R \cosh t_R \quad (4.7)$$

where t_M, x_M are the Minkowski coordinates and r_R, t_R are the Rindler coordinates. Now consider the behavior of the detector as seen in these Rindler coordinates. The space near the horizon looks very hot; it is full of Rindler quanta. Would these quanta ‘burn’ the infalling detector?

At first one may think that an inertial detector in Minkowski space should see nothing. But we have already noted above that the limits placed on the measuring time causes the detector to be excited by vacuum fluctuations. We will now see that such an excitation is of the same kind as that expected in fig.3(b), where we have ‘real’ quanta being radiated at the Rindler temperature by a surface placed just outside the Rindler horizon.

Let the quanta being detected correspond to a scalar field ϕ , which is taken to be in the Minkowski vacuum state $|0\rangle_M$. Since our observations are confined to the right Rindler wedge, we can use the expansion of the field operator in Rindler modes

$$\hat{\phi} = \sum_{\omega} [f_{\omega}(r_R) e^{-i\omega t_R} \hat{a}_{\omega} + f_{\omega}^*(r_R) e^{i\omega t_R} \hat{a}_{\omega}^{\dagger}] \quad (4.8)$$

Let the detector be a 2-level system. We will take it to start in the unexcited state $|i\rangle$, and interactions with ϕ can move it to the state $|f\rangle$. The interaction is described by $\int d\tau \hat{H}_{int}(\tau)$ where

$$\hat{H}_{int}(\tau) = q h(\tau) \hat{O}(\tau) \hat{\phi}(t_R(\tau), r_R(\tau)) \quad (4.9)$$

Here \hat{O} is an operator made out of the detector variables, q is a coupling constant and $0 \leq h(\tau) \leq 1$ is a ‘switching function’ that allows us to switch on and switch off the interaction of the detector with the scalar field ϕ .

The Minkowski vacuum $|0\rangle_M$ can be written in terms of Rindler states of the left (L) and right (R) wedges

$$|0\rangle_M = C \sum_k e^{-\frac{E_k}{2}} |E_k\rangle_L |E_k\rangle_R, \quad C = \left(\sum_k e^{-E_k} \right)^{-\frac{1}{2}} \quad (4.10)$$

Now suppose the interaction is switched on for a brief period as indicated in fig.3(a). Before the interaction is switched on, the state of the overall system is

$$|\Psi\rangle_i = |i\rangle \otimes C \sum_k e^{-\frac{E_k}{2}} |E_k\rangle_L |E_k\rangle_R \quad (4.11)$$

Using first order perturbation theory in the strength of the interaction q , we ask for the amplitude for the transition

$$|i\rangle \otimes |E_k\rangle_L |E_k\rangle_R \rightarrow |f\rangle \otimes |E_k\rangle_L |E_{k'}\rangle_R \quad (4.12)$$

This amplitude is

$$\mathcal{A}_{kk'} = -i \int_{-\infty}^{\infty} d\tau h(\tau) \langle f | \hat{O} | i \rangle_R \langle E_{k'} | \hat{\phi}(t_R(\tau), r_R(\tau)) | E_k \rangle_R \quad (4.13)$$

The quantity ${}_R \langle E_{k'} | \hat{\phi}(t_R(\tau), r_R(\tau)) | E_k \rangle_R$ can be easily computed by writing $|E_k\rangle_R$ in terms of the occupation numbers for different Rindler modes and using the field expansion (4.10). Note that $h(\tau)$ is nonzero only over the part of the detector trajectory indicated in fig.3(a).

The probability for the detector to get excited $|i\rangle \rightarrow |f\rangle$ is then

$$P_{Minkowski} = |C|^2 \sum_k e^{-E_k} \sum_{k'} |\mathcal{A}_{kk'}|^2 \quad (4.14)$$

where the subscript on P indicates that this computation was performed for the Minkowski vacuum situation of fig.3(a). Here the factor e^{-E_k} reflects the fact that the probability of finding the state $|E_k\rangle_R$ in the state (4.10) is

$$p_{E_k} = |C|^2 e^{-E_k} \quad (4.15)$$

Now consider a state that describes a warm body at the same temperature as Rindler space, as shown in fig.3(b). In terms of Rindler eigenstates, this state has a form

$$|\Psi\rangle = \sum_k C_k |E_k\rangle \quad (4.16)$$

Different microstates of the warm body have different coefficients C_k , but the ensemble average over possible microstates will have

$$\langle |C_k|^2 \rangle = |C|^2 e^{-E_k} \quad (4.17)$$

in agreement with (4.15). We again consider the infalling detector with the same interaction (4.9). With the state (4.16) the probability for the detector to get excited is

$$P_{microstate} = \sum_k \sum_{k'} |C_k|^2 |\mathcal{A}_{kk'}|^2 \quad (4.18)$$

Using (4.17) we find that the the ensemble average of the excitation probability for radiation from ‘warm bodies’ is the same as the excitation probability in the Minkowski vacuum when the detection range is confined to be outside the horizon

$$\langle P_{microstate} \rangle = P_{Minkowski} \quad (4.19)$$

In particular, if the infalling body is macroscopic so that it ‘measures’ a large number of quanta, then the effect of radiation in any one microstate will be approximately the same as the effect of vacuum fluctuations when we restrict to the part of the observer worldline that is outside the horizon:

$$P_{microstate} \approx P_{Minkowski} \quad (4.20)$$

A similar effect is also obtained when we consider a detector that has fallen in from near infinity. Quanta at infinity with wavelength λ are wavepackets that have a transverse size $\Delta \gtrsim \lambda$; this is necessary since otherwise the uncertainty principle will give the quantum more transverse momentum $\sim 1/\Delta$ than radial momentum, and the quantum will not really be headed towards the black hole. As the quantum comes closer to the horizon, the wavelength in the radial directions becomes small by blue-shifting, while the transverse size Δ remains unaffected. Thus all quanta falling in from infinity are ‘flattened’ near the horizon. The largeness of Δ compared to the radial wavelengths of Hawking quanta near the horizon means that several Hawking quanta at different angular positions along the horizon can interact with the infalling quantum. Thus we are again let to compute statistical averages, getting a result like (4.20).

4.3 Summary

To summarize, we have compared measurements made by an infalling detector in the case of Minkowski space (fig.3(a)) and in the case of a warm body at the same temperature (fig.3(b)). These two cases are equivalent to the traditional black hole and to a black object with a radiating surface just outside the horizon. While one might at first think that the detector would measure very different things in the two cases, we find that the detector excitation probabilities are actually *similar*. The underlying reason for the similarity is the fact that we need the detection to be completed before the detector reaches the horizon, and this causes vacuum fluctuation excitations in the Minkowski space case that resemble the ‘real’ quanta picked up in the warm body case.

While fuzzballs radiate at exactly the rate expected for Hawking emission, one may envisage a theory other than string theory where the quanta are emitted with energy

$$E \gg kT \tag{4.21}$$

with T the Hawking temperature. In other words, we may give up the thermal spectrum of emission, and have the situation pictured in fig.3(c) where the emitted quantum has wavelength $\lambda \ll d$ at a distance d from the horizon. In this case it *is* possible to make a reliable measurement of the quantum, since ample time is available before the detector reaches the black hole surface. But in this case the emitted radiation will not be carry away the information of the black hole. This follows because the entropy of Hawking radiation (at temperature T) is just ~ 1.3 times the Bekenstein entropy S_{bek} [32]. Taking $E \gg kT$ will give us $N \ll S_{bek}$ quanta to carry out the S_{bek} bits in the black hole, and this is not possible since each quantum carries ~ 1 bit of information.

5 The AMPS argument

In this section we examine the main argument of Almheiri, Marolf, Polchinski and Sully (AMPS). We will note that the measurement they envisage cannot be performed reliably in the given situation, and further, that any conclusions about complementarity cannot be drawn from such a situation.

In outline, the AMPS argument goes as follows:

- (i) If Hawking evaporation is unitary, then the state near the horizon is not the vacuum in an infalling observer’s frame, but involves high-energy excitations.
- (ii) If there are high-energy excitations near the horizon, then an infalling observer will measure physical high energy quanta emerging from the black hole, and get burnt.

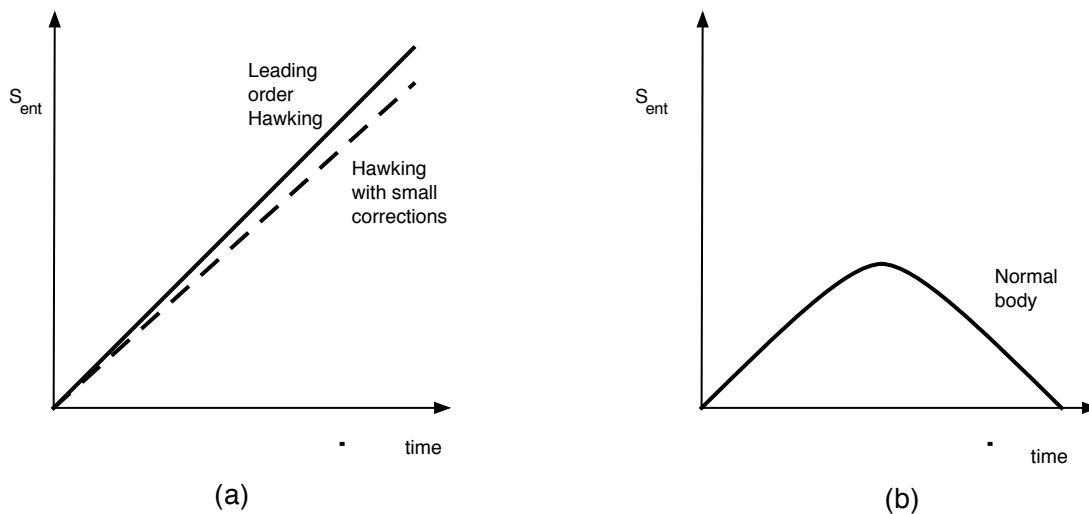


Figure 4: (a) The growth of entanglement entropy for the traditional black hole in the leading order Hawking computation (solid line), and with small corrections allowed (dashed line). (b) The entanglement entropy expected for a normal body [33]; S_{ent} must return to zero when the body radiates away completely.

- (iii) If the observer gets burnt, then we cannot have any complementary description where he falls through without noticing anything at the horizon.

We examine each of these steps in turn.

(i) The need for large corrections at the horizon

In [12] it was shown, using strong subadditivity, that semiclassical physics at the horizon cannot lead to the behavior of entanglement entropy S_{ent} that is expected for normal bodies [33]. The behavior for S_{int} is depicted in fig.4. AMPS try to summarize a version of this argument, but miss a crucial step. We would like to clarify this point since it is important, before continuing with the AMPS argument.

Consider the Hawking pair (2.2) produced in the leading order Hawking process; let the outer and inner members of this pair be called B, C respectively. AMPS consider this leading order process, a fact which is implicit in their assumption that $S_{BC} = 0$; i.e., the produced pair is not entangled with anything else (fig.5(a)). They then use strong subadditivity to argue that S_{ent} cannot return to zero like it should for normal bodies. But this situation does not need the powerful relation of strong subadditivity. In the leading order Hawking process the relation (2.2) tells us that the state of the created pairs is a tensor product of individual pairs (eq. (17) of [12]), and so $S_{ent} = N \ln 2$ after N pairs have been produced. This gives the linearly increasing graph of fig.4(a), and we

do not need strong subadditivity to prove that S_{ent} does not return to zero.

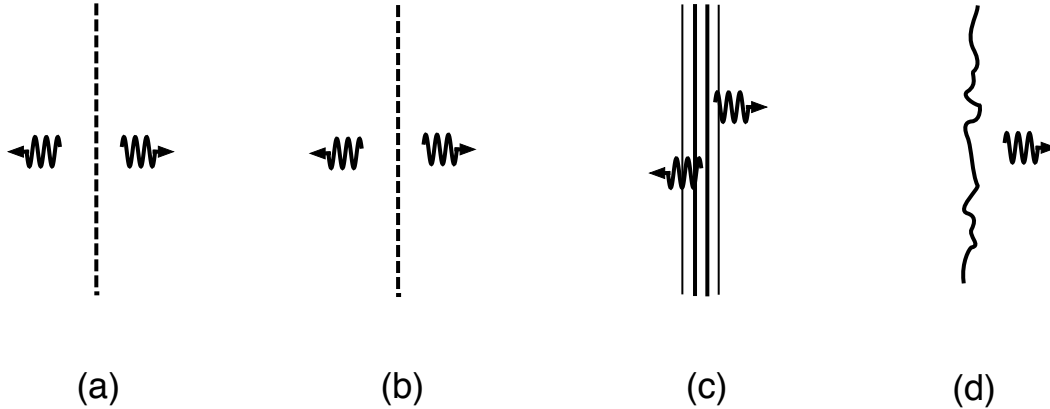


Figure 5: (a) Creation of entangled pairs in the leading order Hawking computation. (b) Small corrections; if these could reproduce the graph fig.4(b) then we would not need a firewall. (c) A firewall that one can pass through; now one can detect the quanta near the horizon. (d) In a fuzzball spacetime ends before the horizon. Hawking radiation is an integral part of the dynamics of the fuzzball.

The important issue, as discussed in Section 2, is whether *subleading* corrections to the leading order Hawking process can make S_{ent} reproduce the behavior of a normal body.⁶ If small corrections could do the job, then we *cannot* conclude that there would be a firewall; we depict this in fig.5(b). To analyze small corrections we have to start from $S_{BC} = \epsilon$ and then we do use strong subadditivity to establish the required inequality (2.4).

To summarize, a smooth vacuum at the horizon leads to the creation of Hawking pairs (2.2), and with (2.4) we see that we cannot get information out in Hawking radiation. Thus if we do wish to have the radiation be unitary, then we must alter the structure of the modes involved in the Hawking process. One may try to restrict the required change to just these modes; this requires us to invoke as yet undiscovered nonlocal effects [15]. If we choose to not do this, then we have an alteration of the physics for *all modes* at the horizon. AMPS take the latter route⁷, and then consider an experiment: they let an infalling observer fall into such a hole and argue he will get ‘burnt’ by the altered structure at the horizon. Further, they argue that getting burnt in this way precludes the possibility of complementarity. We now examine each of these issues in turn.

⁶The possibility that this might happen was raised in [31]. Hawking’s reversal of his belief that information is lost was also implicitly based on the assumption that exponentially small corrections to the leading order process would produce an unentangled state [34].

⁷They consider the possibility of nonlocal effects in a separate discussion later.

(ii) Getting burnt by Hawking quanta

Here AMPS wish to distinguish the traditional black hole from a body that radiates at the Hawking temperature from a surface just outside the horizon. They argue that in the case of the radiating body an infalling observer will observe high energy quanta, while there will be no such quanta observed for the traditional black hole. Let us see what questions we can ask:

(a) The temperature of the radiation is $T \sim m_p$ at a distance l_p from the horizon. If we wish to avoid planck scale physics, then we can try to focus on the radiation a distance d from the horizon with

$$l_p \ll d \ll 2M \quad (5.1)$$

For concreteness, let us think of $d \sim 10^6 l_p$, where we expect the temperature to be high enough to ‘burn’ but the physics is still not the unknown physics at planck scale. For instance, in fig.1 of AMPS, we can ask if we get burnt during the part of the infalling detector where it crosses the shaded region representing the outgoing Hawking quantum. To restrict our question to the required region, we consider the effect of the radiation on a detector which is switched off when we get to a distance closer than $\sim 10^6 l_p$ from the horizon. But in such a situation we have noted in Section 4.2 that the excitation probability of the detector in the case of the traditional horizon and in the case of the firewall are statistically the *same* (eq.(4.20)). Thus we cannot say that we will get burnt in one case and not the other.

(b) The above equality of excitation probabilities resulted from the fact that we were allowed a limited time to make the detection; thus vacuum fluctuations excited the detector even for the traditional hole. We could allow ourselves a longer time for detection if we assumed that we could pass through the firewall to the other side of the horizon, and again find ourselves in a region of low temperature. Such a possibility is pictured in fig.5(b), and in this case we would excite the detector for the firewall but not for the traditional hole. But to have this ‘other side’ to the firewall we need to pass through a ‘wall’ of of planck scale physics. Since the strength of gravitational interactions increases with energy, we can expect that the largest interactions would be when the detector is crossing the region of planck temperature, and so we cannot focus on the issue of detection of quanta with wavelength $\sim 10^6 l_p$ without asking if the theory allows us to pass through the planck temperature region. (In particular, it is hard to imagine a physical model reproducing fig.5(c) which has smooth space on both sides of a planck energy region.) As we note below, the fuzzball microstates of string theory do *not* allow us to pass the planck region; spacetime ends there in a stringy mess. Further, as we will note in part (iii) below, interaction with the planck degrees of freedom is not what precludes the kind of complementarity that we find with fuzzballs; instead, it is

this interaction which transfers information to the collective modes of the fuzzball and leads to a complementary description.

(c) In fig.5(c) we depict the situation with fuzzballs.. The incoming quanta cannot pass through the fuzzball surface, and transfer their energy to excitations of the fuzzball. The fuzzball details and the radiation it emits are parts of the same structure: the radiation is the small time dependent part of the gravitational solution away from $r \approx 2M$. The response of the planck degrees of freedom is encoded in the response (3.5) of the fuzzball, and this effect is expected to dominate over interactions with the radiation tail.

One thing is important to note about this interaction. Let the infalling observer be made of degrees of freedom that evolve slower than the planck scale. Then the observer does not evolve significantly between the time that its coupling to the radiation becomes significant and the time it reaches the fuzzball boundary. Thus it is not clear what ‘burning’ means in this context. The correct question to focus on is not the evolution of the infalling observer, but rather the evolution (3.5) of the fuzzball degrees of freedom that the observer impacts.

(iii) The possibility of complementarity

Finally, let us address the issue of complementarity. AMPS argue that if the infalling quantum ‘burns up’ then we cannot get a complementary description where ‘nothing happens at the horizon’. This desire to avoid any interaction is suggested by the traditional proposal of complementarity. But as we have seen in Section 3.2, there are difficulties with traditional complementarity, and this is not the kind of complementarity that we have proposed.

Consider first our toy example of AdS/CFT duality. A graviton falling onto a D1D5 brane bound state *did* interact strongly on reaching these branes and broke up into a pair of excitations. Yet there was a complementary description where it passed smoothly into an AdS region. Similarly, $E \gg kT$ gravitons falling onto the fuzzball surface *do* interact strongly with the surface and excite the collective dynamics of the fuzzball degrees of freedom; it is this collective dynamics (3.5) which will have a dual representation where to a first approximation the graviton will appear to fall through a horizon.

The moral we would draw is that AMPS are incorrect in saying that complementarity would be impossible if an object encountered strong interactions near the horizon. The situation is quite the opposite: we need *strong* interactions near the horizon to absorb the energy of the infalling quantum into the black hole’s degrees of freedom to get ‘fuzzball complementarity’. If the absorption leads to an approximately faithful map of the infalling quantum’s Hilbert space into a subspace of the black hole degrees of freedom, then we have the possibility of a complementary description of the infall.

6 Discussion

In this paper we have done two things: we summarized how complementarity is conjectured to work with fuzzballs, and we noted how the AMPS argument fails to address the underlying physics in this conjecture. In the discussion below we will put these two parts together, to see more directly where the AMPS argument goes wrong. In short, we will see that complementarity is a story of *two* descriptions of the physics, while AMPS try to have elements of both descriptions in the *same* setting.

For the discussion below, it is helpful to summarize one version of the AMPS argument as follows:

- (1) Suppose the infalling observer sees nothing around $r = 2M$ in some description.
- (2) Then in this description we have a smooth patch of spacetime around the horizon.
- (3) Evolution of vacuum modes in this smooth patch will lead to an entangled hawking pair, and this will lead to the information problem.

The problem with this argument is that the description in which we have (1) (i.e. smooth spacetime) is not the description in which we have high energy Hawking quanta near $r = 2M$. The description in (1) is useful as an approximation for describing the physics of $E \gg kT$ infalling quanta for short times (order $\sim M$). One cannot use the effective smooth spacetime used in this approximation to describe the entanglement of Hawking pairs over the much longer Hawking evaporation time (order $\sim M^2$); in particular one cannot relate the effective description of (1) to the information problem which needs us to talk about the details of $\sim (M/m_p)^2$ Hawking pairs.

Let us now see in more detail how things actually work:

(a) The microstates of the black hole are fuzzballs, which means that the gravitational solution ends just outside $r = 2M$ when the compact directions pinch off; the structure at this location is a quantum mess of KK monopoles, strings, fluxes, etc. (i.e. the set of allowed sources in string theory).

(b) $E \sim kT$ radiation is emitted from these sources, carrying the information of the microstate. A simple model to keep in mind is the computation of [19], where ergoregions near the fuzzball surface emit quanta by ergoregion emission. From this computation we learn that there is no sharp separation between the radiation and the fuzzball: the gravitational field in the ergoregion is unstable and radiates gravitons. If we follow these emitted gravitons back to their source, then we find more and more nonlinear gravitational physics, culminating in the ‘cap’ where the fuzzball solution ends in KK monopoles etc. Thus whenever we ask if we interact with emitted quanta, we might as

well go all the way and ask if we interact with the fully nonlinear ‘cap’.

(c) We recalled the toy example of AdS/CFT, which has similarities and differences with the black hole case. For now we look at the similarities. Suppose we have a bound state of N D1 and N D5 branes. The infalling quanta of fig.2(a) impacts this collection of branes and transfers its energy into excitations of the branes. Similarly, a quantum falling onto the fuzzball transfers its energy to the string theoretic sources (KK monopoles etc) on the fuzzball surface. In (b) we had noted that the radiation from the fuzzball was just the tail end of the full nonlinear KK monopole ‘cap’, so interactions with radiation near the horizon are included in this description.

(d) But in the AdS/CFT case, there is a second description, that of fig.2(b), where the infalling quantum sails smoothly through into an AdS region. In this description we do not see the D1 and D5 branes as something that can be ‘hit’. In the analogous case of fuzzballs, an infalling object does not see the nonlinear KK monopoles etc. near $r = 2M$, but instead sails through smoothly. In particular, it does not see the ‘tail end’ of the nonlinear structure – the radiation from the fuzzball – as high energy quanta that can be ‘hit’.

(e) To understand how it is possible for the infalling observer to sail through smoothly, consider first the AdS/CFT example. If the D1,D5 branes were ‘inert’; i.e., they did not shift their internal state when the infalling object approached, then there would not be any description where the object ‘sailed through’. But in fact the D1D5 brane bound state has a *vast* space of internal excitations, and this changes the situation: the approach of the infalling object creates excitations in this vast space of possibilities, and the dynamics of these excitations is the dominant physics of the combined branes+object system. It is this dynamics that is described by the smooth infall into AdS space.⁸

(f) The fuzzball has a similarly large phase space of deformations, since the number of fuzzball solutions is $Exp[S_{bek}]$. Now we see the basic element missing from the analysis of AMPS. They ask for the dynamics of the infalling object (what it measures etc.) but they ignore the fact that the much more important dynamics is the change of the state of the *fuzzball*: $\sum_i C_i |F_i\rangle \rightarrow \sum_j C'_j |F'_j\rangle$. This latter dynamics is so dominant that one must consider the infalling object and fuzzball as one unified system and then analyze the dynamics. When the infalling object has energy $E \gg kT$, this dynamics has an approximate description valid for short times (order $\sim M$) that mimics infall through a smooth horizon [26, 13, 27, 28]; this is analogous to how in the AdS/CFT case the object

⁸One often thinks of AdS/CFT duality as saying that the gravity variables in AdS can be re-expressed in terms of gauge theory variables on the boundary of AdS. But the origin of this duality is in the context of absorption by D-branes and black holes, and in that context the natural process to consider is the infall of quanta from infinity onto the branes. The largeness of N , the number of branes, leads to the excitation spectrum of the branes as being very dense, and the effective AdS description emerges.

falls through smoothly into an AdS space.

(g) The approximate nature of the ‘smooth infall’ description is important. Since this description is valid only over a time of order $\sim M$, we cannot use this patch of smooth space to argue that entangled Hawking pairs will be created. There is hardly time to create one pair in such a region. We cannot join together many such patches to argue that we have created many entangled pairs, since the description is only valid for short times and does not accurately track $E \sim kT$ physics. The existence of many entangled pairs would have led to the information problem as discussed in Section 2(b),(c); this problem does not arise here since we cannot study the creation of a large set of such pairs in our approximate ‘smooth infall’ description.

To summarize, the error in the AMPS argument can be seen by considering the infall of an observer into a stack of branes. These branes are in a particular internal state, which can be probed by patient low energy scattering experiments from infinity. But the infalling observer reports none of this structure as he approaches the branes; he feels as if it is falling through empty AdS space. This ‘magical disappearance’ of the branes can be traced to the fact that the branes have a vast set of internal states, and the dominant effect of the approach of the observer is to alter the internal state of the branes. Thus the dynamics of the branes+observer system is governed by the evolution of these newly created excitations, and not by the observer scattering off a fixed state of the branes.

We can now see the fundamental role that the fuzzball construction plays in resolving the puzzles with black holes. If we have the traditional Penrose diagram of the hole, with vacuum at the horizon, then we get the creation of entangled pairs, and we cannot evade the Hawking information loss problem [2, 12]. But in string theory we find that there is very nontrivial structure at the horizon: the KK monopoles etc at the fuzzball surface carry ‘real’ degrees of freedom that radiate unitarily like a normal body. This resolves the information paradox. But we can ask a different question: what happens when a high energy ($E \gg kT$) quantum impacts the fuzzball surface? In this case the physics is analogous to what we find in AdS/CFT: the KK monopole and other string theoretic degrees of freedom on the fuzzball surface act like the branes in the D1D5 system. The infalling observer reports nothing special as it approaches these objects, since the dominant dynamics is that of *exciting* the fuzzball degrees of freedom, not the response of the observer. AMPS implicitly assume that they are falling towards a radiating surface that is *inert* to such excitations, and thus miss the physics of free infall which is common to the fuzzball and the AdS/CFT cases.

In the context of the argument (1)-(3) listed at the start of this section, we see that the implication (1) \rightarrow (2) is misleading. It is not that we don’t have structure at the location of the branes; rather, the infalling observer does not report such structure. The

patch of smooth spacetime in (2) is an effective description of the $\sum_i C_i |F_i\rangle \rightarrow \sum_j C'_j |F'_j\rangle$ dynamics which describes the excitations of the impacted fuzzballs; it is not the actual gravitational solution at the horizon. The implication (2) \rightarrow (3) does not work since this effective description cannot be applied to a region larger than $\sim M$ which would be needed to create a large number of entangled pair. In general there are *two* descriptions involved: (i) the actual microscopic fuzzball which carries all information of the state and radiates unitarily, and (ii) the approximate short time description of collective modes, that mimics free infall. AMPS do not differentiate carefully these two descriptions, and that leads them to an apparent contradiction.

In conclusion, the AMPS argument does not apply to the process by which complementarity is conjectured to arise in the fuzzball picture. But it is a very interesting argument to consider, since it brings out clearly the various important physical principles involved in the quantum dynamics of black holes.

Acknowledgements

This work was supported in part by DOE grant DE-FG02-91ER-40690. We thank the authors of [9] as well as Iosif Bena, Borun Chowdhury, Stefano Giusto, Oleg Lunin, Lenny Susskind and Nick Warner for discussions. In particular we are grateful to Don Marolf for patiently explaining to us the nature of the AMPS argument.

A Timescale for detection

Here we note that a detector needs a proper time $\Delta\tau \gtrsim \lambda$ to detect a quantum of wavelength λ . Since this argument is well known, we will describe it for the simple case of a detector at rest in the Minkowski vacuum; the extension to other situations is straightforward.

We assume for simplicity that the metric is time independent in our choice of coordinates. The field operator can be expanded as

$$\hat{\Phi} = \sum_k \left[\frac{1}{\sqrt{2\omega_k}} e^{i(kx - \omega_k t)} \hat{a}_k + \frac{1}{\sqrt{2\omega_k}} e^{-i(kx - \omega_k t)} \hat{a}_k^\dagger \right], \quad [\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{k,k'} \quad (\text{A.1})$$

We take the detector to be a harmonic oscillator

$$\hat{\Psi} = \frac{1}{\sqrt{2\Omega}} e^{-i\Omega\tau} \hat{A} + \frac{1}{\sqrt{2\Omega}} e^{i\Omega\tau} \hat{A}^\dagger, \quad [\hat{A}, \hat{A}^\dagger] = 1 \quad (\text{A.2})$$

The interaction along the worldline is given by $\int d\tau \hat{H}_{int}(\tau)$ where

$$\hat{H}_{int}(\tau) = g h(\tau) \hat{\Phi}(t(\tau), x(\tau)) \hat{\Psi}(\tau) \quad (\text{A.3})$$

Here q is a coupling constant and $0 \leq h(\tau) \leq 1$ is a function that allows us to switch on and switch off the detector.

We start at $\tau \rightarrow -\infty$ with the detector in the ground state: $\hat{A}|0\rangle_A = 0$. Let us also take the spacetime to be empty of quanta: $\hat{a}_k|0\rangle_a = 0$. We take first order perturbation theory in q . The amplitude to reach the state $|1\rangle_A|1\rangle_k \equiv \hat{A}^\dagger \hat{a}_k^\dagger|0\rangle_A|0\rangle_a$ is

$$\mathcal{A} = -iq \int_{-\infty}^{\infty} d\tau \frac{1}{\sqrt{2\omega_k}} \frac{1}{\sqrt{2\Omega}} h(\tau) e^{i\Omega\tau} e^{-ikx(\tau) + i\omega_k t(\tau)} \quad (\text{A.4})$$

We take

$$h(\tau) = e^{-(\frac{\tau}{\Delta\tau})^2} \quad (\text{A.5})$$

which corresponds to making a measurement over an interval $\sim \Delta\tau$. We also let the detector trajectory to describe a detector at rest at $x = 0$, which gives $x(\tau) = 0, t(\tau) = \tau$ for all τ . This gives

$$\mathcal{A} = -iq \int_{-\infty}^{\infty} d\tau \frac{1}{\sqrt{2\omega_k}} \frac{1}{\sqrt{2\Omega}} e^{-(\frac{\tau}{\Delta\tau})^2} e^{i(\Omega + \omega_k)\tau} = -iq \frac{1}{\sqrt{2\omega_k}} \frac{1}{\sqrt{2\Omega}} \Delta\tau \sqrt{\pi} e^{-\frac{1}{4}(\Delta\tau)^2(\Omega + \omega_k)^2} \quad (\text{A.6})$$

Keeping the detector on for all time is equivalent to taking $\Delta\tau \rightarrow \infty$, in which case we get $\mathcal{A} = 0$. So the detector does not get excited, which is expected since we started with empty Minkowski space.

But now consider a situation where the detector is switched on and off in a comparatively short interval, as would need to be the case if one was trying to detect a Hawking quantum by an infalling detector before the detector hit the black hole surface. For detection times shorter than the wavelengths we want to measure

$$\Delta\tau \lesssim \frac{1}{(\Omega + \omega_k)} \quad (\text{A.7})$$

we get

$$\mathcal{A} \sim -iq \frac{1}{\sqrt{2\omega_k}} \frac{1}{\sqrt{2\Omega}} \Delta\tau \sqrt{\pi} \neq 0 \quad (\text{A.8})$$

so we pick up vacuum fluctuations in the detector.

To summarize, suppose we make a detector with frequency Ω to pick up quanta of wavelength $\lambda \sim \Omega^{-1}$. Then the effect of vacuum fluctuations will be comparable to the effect of ‘real quanta’ if

$$\Delta\tau \lesssim \frac{1}{(\Omega + \omega_k)} < \frac{1}{\Omega} \sim \lambda \quad (\text{A.9})$$

B Wavelength of Hawking quanta

Consider the Schwarzschild black hole

$$ds^2 = -(1 - \frac{2M}{r})dt^2 + \frac{dr^2}{1 - \frac{2M}{r}} + r^2 d\Omega_2^2 \quad (\text{B.1})$$

The temperature is $\frac{1}{8\pi M}$, so the wavelength of Hawking quanta at infinity is $\lambda_\infty \sim M$. The wavelength of such a quantum at any position r is

$$\lambda \sim (-g_{tt})^{\frac{1}{2}} \lambda_\infty \sim M(1 - \frac{2M}{r})^{\frac{1}{2}} \quad (\text{B.2})$$

Near the horizon $(r - 2M) \ll 2M$ we can use Rindler coordinates

$$t_R = \frac{t}{4M}, \quad r_R = \sqrt{8M(r - 2M)} \quad (\text{B.3})$$

This gives the metric in the time and radial directions

$$ds^2 \approx -r_R^2 dt_R^2 + dr_R^2 \quad (\text{B.4})$$

From now on we restrict attention to just these directions. In this near-horizon region we have for the wavelength of radiated quanta

$$\lambda \sim M^{\frac{1}{2}}(r - 2M)^{\frac{1}{2}} \sim r_R \quad (\text{B.5})$$

From (B.4) we see that the distance from the horizon measured on a constant t_R slice is $d = r_R$. Thus if a black hole emits radiation at the Hawking temperature, then the wavelength of these quanta at a distance d from the horizon is

$$\lambda \sim d \quad (\text{B.6})$$

This is the wavelength measured along a slice of constant Schwarzschild time t . If this quantum is encountered by an infalling detector, then the effective wavelength will be Lorentz contracted. Let the proper velocity of the detector in a local Lorentz frame oriented along the Schwarzschild t, r directions, be

$$U^{\hat{t}} = \cosh \alpha, \quad U^{\hat{r}} = -\sinh \alpha \quad (\text{B.7})$$

The momentum vector of an outgoing massless quantum in the local Lorentz frame is

$$(p^{\hat{t}}, p^{\hat{r}}) \sim \left(\frac{1}{\lambda}, \frac{1}{\lambda}\right) \quad (\text{B.8})$$

The energy of the quantum as measured by the detector is then

$$E = -p_\mu U^\mu \sim \frac{1}{\lambda}(\cosh \alpha + \sinh \alpha) = \frac{1}{\lambda}e^\alpha \quad (\text{B.9})$$

and the effective wavelength that is seen by the infalling detector is then

$$\lambda_{eff} \sim \lambda e^{-\alpha} \sim d e^{-\alpha} \quad (\text{B.10})$$

where as above, d is the distance measured from the horizon in the Schwarzschild frame along a $t = \text{const}$ slice.

C Proper time along infalling geodesic

We wish to ask how much proper time $\Delta\tau$ elapses along a geodesic between the time it is at a distance d from $r = 2M$ and the time it hits the black hole surface at $r = 2M$. Since we are working near the horizon, we use the Rindler coordinates (B.3). The Kruskal-type coordinates appropriate to a freely falling observer are given locally by taking the Minkowski coordinates related to t_R, r_R by

$$t_M = r_R \sinh t_R, \quad x_M = r_R \cosh t_R \quad (\text{C.1})$$

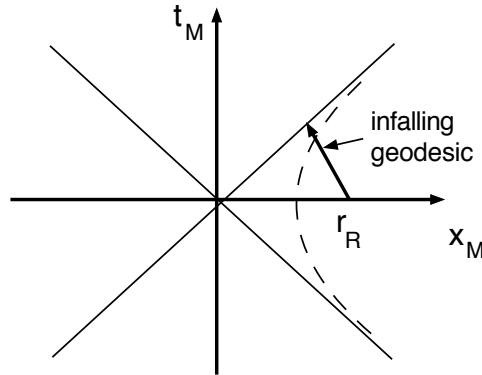


Figure 6: The Rindler coordinates near the horizon, and the corresponding Minkowski coordinates. The infalling geodesic starts at $r_R = d, t_R = 0$ and ends at the Rindler horizon.

Fig.6 shows the geodesic that we follow. This geodesic is a straight line in the local Minkowski coordinates

$$t_M = \cosh \alpha \tau, \quad x_M = -\sinh \alpha \tau + d \quad (\text{C.2})$$

We have taken the geodesic to start with $\tau = 0$ at position $r_R = d$ and time $t_R = 0$. Here α is a constant that gives the velocity of infall; note that it is the same α as the one that appears in (B.7). The geodesic crosses the horizon $t_M = r_M$ at proper time τ_f with

$$\cosh \alpha \tau_f = -\sinh \alpha \tau_f + d, \quad \Rightarrow \quad \tau_f = d e^{-\alpha} \quad (\text{C.3})$$

Thus if an observer on an infalling trajectory tries to detect a quantum at distance d from the horizon, then the time he has available to make the detection is

$$\Delta\tau_{\text{available}} < \tau_f = d e^{-\alpha} \quad (\text{C.4})$$

References

- [1] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973).
- [2] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975) [Erratum-ibid. **46**, 206 (1976)]; S. W. Hawking, Phys. Rev. D **14**, 2460 (1976).
- [3] G. 't Hooft, hep-th/0003004; L. Susskind, J. Math. Phys. **36**, 6377 (1995) [arXiv:hep-th/9409089]; L. Susskind, L. Thorlacius, J. Uglum, Phys. Rev. **D48**, 3743-3761 (1993). [hep-th/9306069]; L. Susskind, Phys. Rev. Lett. **71**, 2367-2368 (1993). [hep-th/9307168]; D. A. Lowe, J. Polchinski, L. Susskind *et al.*, Phys. Rev. **D52**, 6997-7010 (1995). [hep-th/9506138].
- [4] A. Strominger and C. Vafa, Phys. Lett. B **379**, 99 (1996) [hep-th/9601029].
- [5] O. Lunin and S. D. Mathur, Nucl. Phys. B **623**, 342 (2002) [arXiv:hep-th/0109154].
- [6] O. Lunin and S. D. Mathur, Phys. Rev. Lett. **88**, 211303 (2002) [arXiv:hep-th/0202072]; O. Lunin, J. M. Maldacena and L. Maoz, arXiv:hep-th/0212210; S. D. Mathur, A. Saxena and Y. K. Srivastava, Nucl. Phys. B **680**, 415 (2004) [arXiv:hep-th/0311092]; O. Lunin, JHEP **0404**, 054 (2004) [arXiv:hep-th/0404006]; S. Giusto, S. D. Mathur and A. Saxena, Nucl. Phys. B **710**, 425 (2005) [arXiv:hep-th/0406103].
- [7] I. Bena and N. P. Warner, Adv. Theor. Math. Phys. **9**, 667 (2005) [arXiv:hep-th/0408106]; I. Bena and N. P. Warner, Phys. Rev. D **74**, 066001 (2006) [arXiv:hep-th/0505166]; P. Berglund, E. G. Gimon and T. S. Levi, JHEP **0606**, 007 (2006) [arXiv:hep-th/0505167]; A. Saxena, G. Potvin, S. Giusto and A. W. Peet, JHEP **0604**, 010 (2006) [arXiv:hep-th/0509214]; V. Balasubramanian, E. G. Gimon and T. S. Levi, JHEP **0801**, 056 (2008) [arXiv:hep-th/0606118]; J. de Boer, F. Denef, S. El-Showk, I. Messamah and D. Van den Bleeken, JHEP **0811**, 050 (2008) [arXiv:0802.2257 [hep-th]]; I. Bena, J. de Boer, M. Shigemori and N. P. Warner, arXiv:1107.2650 [hep-th];
- [8] S. Giusto, R. Russo and D. Turton, JHEP **1111**, 062 (2011) [arXiv:1108.6331 [hep-th]]; S. D. Mathur and D. Turton, arXiv:1112.6413 [hep-th]; S. D. Mathur and D. Turton, arXiv:1202.6421 [hep-th].
- [9] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, arXiv:1207.3123 [hep-th].
- [10] R. Bousso, arXiv:1207.5192 [hep-th]. D. Harlow, arXiv:1207.6243 [hep-th]. Y. Nomura, J. Varela and S. J. Weinberg, arXiv:1207.6626 [hep-th].

- [11] S. D. Mathur, Fortsch. Phys. **53**, 793 (2005) [hep-th/0502050]; I. Bena and N. P. Warner, Lect. Notes Phys. **755**, 1 (2008) [arXiv:hep-th/0701216]; K. Skenderis and M. Taylor, Phys. Rept. **467**, 117 (2008) [arXiv:0804.0552 [hep-th]]; S. D. Mathur, arXiv:0810.4525 [hep-th]; B. D. Chowdhury and A. Virmani, arXiv:1001.1444 [hep-th].
- [12] S. D. Mathur, Class. Quant. Grav. **26**, 224001 (2009) [arXiv:0909.1038 [hep-th]].
- [13] S. D. Mathur and C. J. Plumberg, JHEP **1109**, 093 (2011) [arXiv:1101.4899 [hep-th]].
- [14] B. Czech, K. Larjo and M. Rozali, “Black Holes as Rubik’s Cubes,” JHEP **1108**, 143 (2011) [arXiv:1106.5229 [hep-th]]; S. G. Avery, “Qubit Models of Black Hole Evaporation,” arXiv:1109.2911 [hep-th].
- [15] S. B. Giddings, “Models for unitary black hole disintegration,” Phys. Rev. D **85**, 044038 (2012) [arXiv:1108.2015 [hep-th]]; S. B. Giddings, “Black holes, quantum information, and unitary evolution,” arXiv:1201.1037 [hep-th]; S. B. Giddings and Y. Shi, “Quantum information transfer and models for black hole mechanics,” arXiv:1205.4732 [hep-th].
- [16] S. D. Mathur, arXiv:1201.2079 [hep-th].
- [17] S. D. Mathur, Nucl. Phys. B **529**, 295 (1998) [arXiv:hep-th/9706151].
- [18] S. R. Das and S. D. Mathur, Phys. Lett. B **375**, 103 (1996) [arXiv:hep-th/9601152].
- [19] V. Jejjala, O. Madden, S. F. Ross and G. Titchener, Phys. Rev. D **71**, 124030 (2005) [arXiv:hep-th/0504181]. V. Cardoso, O. J. C. Dias, J. L. Hovdebo and R. C. Myers, Phys. Rev. D **73**, 064031 (2006) [arXiv:hep-th/0512277]; B. D. Chowdhury and S. D. Mathur, Class. Quant. Grav. **25**, 135005 (2008) [arXiv:0711.4817 [hep-th]]; B. D. Chowdhury and S. D. Mathur, Class. Quant. Grav. **25**, 225021 (2008) [arXiv:0806.2309 [hep-th]]; B. D. Chowdhury and S. D. Mathur, Class. Quant. Grav. **26**, 035006 (2009) [arXiv:0810.2951 [hep-th]].
- [20] S. D. Mathur, arXiv:0805.3716 [hep-th].
- [21] S. D. Mathur, Int. J. Mod. Phys. D **18**, 2215 (2009) [arXiv:0905.4483 [hep-th]].
- [22] O. Lunin and S. D. Mathur, Nucl. Phys. B **642**, 91 (2002) [hep-th/0206107].
- [23] J. M. Maldacena, Adv. Theor. Math. Phys. **2**, 231 (1998) [Int. J. Theor. Phys. **38**, 1113 (1999)] [arXiv:hep-th/9711200]; E. Witten, Adv. Theor. Math. Phys. **2**, 253

- (1998) [hep-th/9802150]; S. S. Gubser, I. R. Klebanov and A. M. Polyakov, Phys. Lett. B **428**, 105 (1998) [hep-th/9802109].
- [24] C. G. Callan and J. M. Maldacena, Nucl. Phys. B **472**, 591 (1996) [arXiv:hep-th/9602043]; A. Dhar, G. Mandal and S. R. Wadia, Phys. Lett. B **388**, 51 (1996) [arXiv:hep-th/9605234]. S. R. Das and S. D. Mathur, Nucl. Phys. B **478**, 561 (1996) [arXiv:hep-th/9606185]; S. R. Das and S. D. Mathur, Nucl. Phys. B **482**, 153 (1996) [arXiv:hep-th/9607149]; J. M. Maldacena and A. Strominger, Phys. Rev. D **55**, 861 (1997) [arXiv:hep-th/9609026].
- [25] S. R. Das and S. D. Mathur, Nucl. Phys. B **482**, 153 (1996) [hep-th/9607149].
- [26] S. D. Mathur, Class. Quant. Grav. **28**, 125010 (2011) [arXiv:1012.2101 [hep-th]].
- [27] S. D. Mathur, arXiv:1205.0776 [hep-th].
- [28] S. D. Mathur, arXiv:1207.5431 [hep-th].
- [29] M. Van Raamsdonk, arXiv:0907.2939 [hep-th]; M. Van Raamsdonk, Gen. Rel. Grav. **42**, 2323 (2010) [arXiv:1005.3035 [hep-th]].
- [30] W. Israel, Phys. Lett. A **57**, 107 (1976).
- [31] J. M. Maldacena, JHEP **0304**, 021 (2003) [arXiv:hep-th/0106112].
- [32] W. H. Zurek, Phys. Rev. Lett. **49**, 1683 (1982).
- [33] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993) [arXiv:gr-qc/9305007].
- [34] S. W. Hawking, Phys. Rev. D **72**, 084013 (2005) [arXiv:hep-th/0507171].